A company has recently created a new department to handle their services workload. An IT team has been asked to create a custom VPC to isolate the resources created in this new department. They have set up the public subnet and internet gateway (IGW). However, they are not able to ping the Amazon EC2 instances with elastic IP address (EIP) launched in the newly created VPC.

As a Solutions Architect, the team has requested your help. How will you troubleshoot this scenario? (Select two)

**Check if the route table is configured with internet gateway**

An internet gateway (IGW) is a horizontally scaled, redundant, and highly available VPC component that allows communication between instances in your VPC and the internet. An internet gateway serves two purposes: to provide a target in your VPC route tables for internet-routable traffic, and to perform network address translation (NAT) for instances that have been assigned public IPv4 addresses. An internet gateway supports IPv4 and IPv6 traffic.

To enable access to or from the internet for instances in a subnet in a VPC, you must do the following: 1. Attach an internet gateway to your VPC. 2. Add a route to your subnet's route table that directs internet-bound traffic to the internet gateway. 3. Ensure that instances in your subnet have a globally unique IP address 4. Ensure that your network access control lists and security group rules allow the relevant traffic to flow to and from your instance.

A route table contains a set of rules, called routes, that are used to determine where network traffic from your subnet or gateway is directed. After creating an IGW, make sure the route tables are updated. Additionally, ensure the security group allows the ICMP protocol for ping requests.

**Check if the security groups allow ping from the source**

A security group acts as a virtual firewall that controls the traffic for one or more instances. When you launch an instance, you can specify one or more security groups; otherwise, AWS uses the default security group. You can add rules to each security group that allow traffic to or from its associated instances. You can modify the rules for a security group at any time; the new rules are automatically applied to all instances that are associated with the security group. To decide whether to allow traffic to reach an instance, all the rules from all the security groups that are associated with the instance are evaluated.

The following are the characteristics of security group rules: 1. By default, security groups allow all outbound traffic. 2. Security group rules are always permissive; you can't create rules that deny access. 3. Security groups are stateful

The engineering team at an e-commerce company has been tasked with migrating to a serverless architecture. The team wants to focus on the key points of consideration when using AWS Lambda as a backbone for this architecture.

As a Solutions Architect, which of the following options would you identify as correct for the given requirement? (Select three)

* **The bigger your deployment package, the slower your AWS Lambda function will cold-start. Hence, AWS suggests packaging dependencies as a separate package from the actual AWS Lambda package**

**(Incorrect)**

* **If you intend to reuse code in more than one AWS Lambda function, you should consider creating an AWS Lambda Layer for the reusable code**

**(Correct)**

* **By default, AWS Lambda functions always operate from an AWS-owned VPC and hence have access to any public internet address or public AWS APIs. Once an AWS Lambda function is VPC-enabled, it will need a route through a Network Address Translation gateway (NAT gateway) in a public subnet to access public resources**

**(Correct)**

* **Since AWS Lambda functions can scale extremely quickly, it's a good idea to deploy a Amazon CloudWatch Alarm that notifies your team when function metrics such as ConcurrentExecutions or Invocations exceeds the expected threshold**

**(Correct)**

* **Serverless architecture and containers complement each other but you cannot package and deploy AWS Lambda functions as container images**
* **AWS Lambda allocates compute power in proportion to the memory you allocate to your function. AWS, thus recommends to over provision your function time out settings for the proper performance of AWS Lambda functions**
* An e-commerce company wants to migrate its on-premises application to AWS. The application consists of application servers and a Microsoft SQL Server database. The solution should result in the maximum possible availability for the database layer while minimizing operational and management overhead.
* As a solutions architect, which of the following would you recommend to meet the given requirements?
* **Migrate the data to Amazon RDS for SQL Server database in a Multi-AZ deployment**
* Amazon RDS supports Multi-AZ deployments for Microsoft SQL Server by using either SQL Server Database Mirroring (DBM) or Always On Availability Groups (AGs). Amazon RDS monitors and maintains the health of your Multi-AZ deployment. If problems occur, Amazon RDS automatically repairs unhealthy database instances, reestablishes synchronization, and initiates failovers.
* Multi-AZ deployments provide increased availability, data durability, and fault tolerance for database instances. In the event of planned database maintenance or unplanned service disruption, Amazon RDS automatically fails over to the up-to-date secondary database instance. This functionality lets database operations resume quickly without manual intervention. The primary and standby instances use the same endpoint, whose physical network address transitions to the secondary replica as part of the failover process. You don't have to reconfigure your application when a failover occurs.
* This option provides the maximum possible availability for the database layer while minimizing operational and management overhead.

A company wants to grant access to an Amazon S3 bucket to users in its own AWS account as well as to users in another AWS account. Which of the following options can be used to meet this requirement?

**Use a bucket policy to grant permission to users in its account as well as to users in another account**

A bucket policy is a type of resource-based policy that can be used to grant permissions to the principal that is specified in the policy. Principals can be in the same account as the resource or in other accounts. For cross-account permissions to other AWS accounts or users in another account, you must use a bucket policy.

A company wants to adopt a hybrid cloud infrastructure where it uses some AWS services such as Amazon S3 alongside its on-premises data center. The company wants a dedicated private connection between the on-premise data center and AWS. In case of failures though, the company needs to guarantee uptime and is willing to use the public internet for an encrypted connection.

What do you recommend? (Select two)

**Use AWS Direct Connect connection as a primary connection**

AWS Direct Connect lets you establish a dedicated network connection between your network and one of the AWS Direct Connect locations. Using industry-standard 802.1q VLANs, this dedicated connection can be partitioned into multiple virtual interfaces. AWS Direct Connect does not involve the Internet; instead, it uses dedicated, private network connections between your intranet and Amazon VPC.

**Use AWS Site-to-Site VPN as a backup connection**

AWS Site-to-Site VPN enables you to securely connect your on-premises network or branch office site to your Amazon Virtual Private Cloud (Amazon VPC). You can securely extend your data center or branch office network to the cloud with an AWS Site-to-Site VPN connection. A VPC VPN Connection utilizes IPSec to establish encrypted network connectivity between your intranet and Amazon VPC over the Internet. VPN Connections can be configured in minutes and are a good solution if you have an immediate need, have low to modest bandwidth requirements, and can tolerate the inherent variability in Internet-based connectivity.

AWS Direct Connect as a primary connection guarantees great performance and security (as the connection is private). Using Direct Connect as a backup solution would work but probably carries a risk it would fail as well. As we don't mind going over the public internet (which is reliable, but less secure as connections are going over the public route), we should use a Site to Site VPN which offers an encrypted connection to handle failover scenarios.

You have an Amazon S3 bucket that contains files in two different folders - s3://my-bucket/images and s3://my-bucket/thumbnails. When an image is first uploaded and new, it is viewed several times. But after 45 days, analytics prove that image files are on average rarely requested, but the thumbnails still are. After 180 days, you would like to archive the image files and the thumbnails. Overall you would like the solution to remain highly available to prevent disasters happening against a whole Availability Zone (AZ).

How can you implement an efficient cost strategy for your Amazon S3 bucket? (Select two)

* **Create a Lifecycle Policy to transition all objects to Amazon S3 Standard IA after 45 days**
* **Create a Lifecycle Policy to transition all objects to Amazon S3 Glacier after 180 days**

**(Correct)**

* **Create a Lifecycle Policy to transition objects to Amazon S3 One Zone IA using a prefix after 45 days**

**(Incorrect)**

* **Create a Lifecycle Policy to transition objects to Amazon S3 Standard IA using a prefix after 45 days**

**(Correct)**

* **Create a Lifecycle Policy to transition objects to Amazon S3 Glacier using a prefix after 180 days**

**(Incorrect)**

A leading e-commerce company runs its IT infrastructure on AWS Cloud. The company has a batch job running at 7AM daily on an Amazon RDS database. It processes shipping orders for the past day, and usually gets around 2000 records that need to be processed sequentially in a batch job via a shell script. The processing of each record takes about 3 seconds.

What platform do you recommend to run this batch job?

**Amazon Elastic Compute Cloud (Amazon EC2)**

Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides secure, resizable compute capacity in the cloud. It is designed to make web-scale cloud computing easier for developers. Amazon EC2’s simple web service interface allows you to obtain and configure capacity with minimal friction. It provides you with complete control of your computing resources and lets you run on Amazon’s proven computing environment. AWS Batch can be used to plan, schedule, and execute your batch computing workloads on Amazon EC2 Instances. Amazon EC2 is the right choice as it can accommodate batch processing and run customized scripts, as is the needed requirement.

You started a new job as a solutions architect at a company that has both AWS experts and people learning AWS. Recently, a developer misconfigured a newly created Amazon RDS database which resulted in a production outage.

How can you ensure that Amazon RDS specific best practices are incorporated into a reusable infrastructure template to be used by all your AWS users?

* **Attach an IAM policy to interns preventing them from creating an Amazon RDS database**

**(Incorrect)**

* **Store your recommendations in a custom AWS Trusted Advisor rule**
* **Use AWS CloudFormation to manage Amazon RDS databases**

**(Correct)**

* **Create an AWS Lambda function which sends emails when it finds misconfigured Amazon RDS databases**
* **Use AWS CloudFormation to manage Amazon RDS databases**
* AWS CloudFormation provides a common language for you to model and provision AWS and third-party application resources in your cloud environment. AWS CloudFormation allows you to use programming languages or a simple text file to model and provision, in an automated and secure manner, all the resources needed for your applications across all regions and accounts. This gives you a single source of truth for your AWS and third-party resources.
* AWS CloudFormation allows you to keep your infrastructure as code and re-use the best practices around your company for configuration parameters. Therefore, this is the correct option for the given use-case.

A developer in your company has set up a classic 2 tier architecture consisting of an Application Load Balancer and an Auto Scaling group (ASG) managing a fleet of Amazon EC2 instances. The Application Load Balancer is deployed in a subnet of size 10.0.1.0/24 and the Auto Scaling group is deployed in a subnet of size 10.0.4.0/22.

As a solutions architect, you would like to adhere to the security pillar of the well-architected framework. How do you configure the security group of the Amazon EC2 instances to only allow traffic coming from the Application Load Balancer?

* **Add a rule to authorize the security group of the Auto Scaling group**
* **Add a rule to authorize the CIDR 10.0.1.0/24**

**(Incorrect)**

* **Add a rule to authorize the security group of the Application Load Balancer**

**(Correct)**

* **Add a rule to authorize the CIDR 10.0.4.0/22**
* You are working as a Solutions Architect for a photo processing company that has a proprietary algorithm to compress an image without any loss in quality. Because of the efficiency of the algorithm, your clients are willing to wait for a response that carries their compressed images back. You also want to process these jobs asynchronously and scale quickly, to cater to the high demand. Additionally, you also want the job to be retried in case of failures.
* Which combination of choices do you recommend to minimize cost and comply with the requirements? (Select two)
* **Amazon EC2 Spot Instances**
* A Spot Instance is an unused Amazon EC2 instance that is available for less than the On-Demand price. Because Spot Instances enable you to request unused Amazon EC2 instances at steep discounts, you can lower your Amazon EC2 costs significantly. The hourly price for a Spot Instance is called a Spot price. The Spot price of each instance type in each Availability Zone (AZ) is set by Amazon EC2 and adjusted gradually based on the long-term supply of and demand for Spot Instances. Your Spot Instance runs whenever capacity is available and the maximum price per hour for your request exceeds the Spot price.
* To process these jobs, due to the unpredictable nature of their volume, and the desire to save on costs, spot Instances are recommended as compared to on-demand instances. As spot instances are cheaper than reserved instances and do not require long term commitment, spot instances are a better fit for the given use-case.
* **Amazon Simple Queue Service (Amazon SQS)**
* Amazon Simple Queue Service (Amazon SQS) is a fully managed message queuing service that enables you to decouple and scale microservices, distributed systems, and serverless applications. SQS offers two types of message queues. Standard queues offer maximum throughput, best-effort ordering, and at-least-once delivery. Amazon SQS FIFO (First-In-First-out) queues are designed to guarantee that messages are processed exactly once, in the exact order that they are sent.
* Amazon SQS will allow you to buffer the image compression requests and process them asynchronously. It also has a direct built-in mechanism for retries and scales seamlessly.
* A healthcare company is evaluating storage options on Amazon S3 to meet regulatory guidelines. The data should be stored in such a way on Amazon S3 that it cannot be deleted until the regulatory time period has expired.
* As a solutions architect, which of the following would you recommend for the given requirement?
* **Use Amazon S3 Object Lock**
* Amazon S3 Object Lock is an Amazon S3 feature that allows you to store objects using a write once, read many (WORM) model. You can use WORM protection for scenarios where it is imperative that data is not changed or deleted after it has been written. Whether your business has a requirement to satisfy compliance regulations in the financial or healthcare sector, or you simply want to capture a golden copy of business records for later auditing and reconciliation, Amazon S3 Object Lock is the right tool for you. Object Lock can help prevent objects from being deleted or overwritten for a fixed amount of time or indefinitely.
* The engineering team at a leading e-commerce company is anticipating a surge in the traffic because of a flash sale planned for the weekend. You have estimated the web traffic to be 10x. The content of your website is highly dynamic and changes very often.
* As a Solutions Architect, which of the following options would you recommend to make sure your infrastructure scales for that day?
* **Use an Auto Scaling Group**
* An Auto Scaling group (ASG) contains a collection of Amazon EC2 instances that are treated as a logical grouping for automatic scaling and management. An Auto Scaling group also enables you to use Amazon EC2 Auto Scaling features such as health check replacements and scaling policies. Both maintaining the number of instances in an Auto Scaling group and automatic scaling are the core functionality of the Amazon EC2 Auto Scaling service.
* The size of an Auto Scaling group depends on the number of instances that you set as the desired capacity. You can adjust its size to meet demand, either manually or by using automatic scaling.
* An Auto Scaling group starts by launching enough instances to meet its desired capacity. It maintains this number of instances by performing periodic health checks on the instances in the group. The Auto Scaling group continues to maintain a fixed number of instances even if an instance becomes unhealthy. If an instance becomes unhealthy, the group terminates the unhealthy instance and launches another instance to replace it.
* Auto Scaling group is the correct answer here.
* A Big Data analytics company writes data and log files in Amazon S3 buckets. The company now wants to stream the existing data files as well as any ongoing file updates from Amazon S3 to Amazon Kinesis Data Streams.
* As a Solutions Architect, which of the following would you suggest as the fastest possible way of building a solution for this requirement?
* **Leverage AWS Database Migration Service (AWS DMS) as a bridge between Amazon S3 and Amazon Kinesis Data Streams**
* You can achieve this by using AWS Database Migration Service (AWS DMS). AWS DMS enables you to seamlessly migrate data from supported sources to relational databases, data warehouses, streaming platforms, and other data stores in AWS cloud.
* The given requirement needs the functionality to be implemented in the least possible time. You can use AWS DMS for such data-processing requirements. AWS DMS lets you expand the existing application to stream data from Amazon S3 into Amazon Kinesis Data Streams for real-time analytics without writing and maintaining new code. AWS DMS supports specifying Amazon S3 as the source and streaming services like Kinesis and Amazon Managed Streaming of Kafka (Amazon MSK) as the target. AWS DMS allows migration of full and change data capture (CDC) files to these services. AWS DMS performs this task out of box without any complex configuration or code development. You can also configure an AWS DMS replication instance to scale up or down depending on the workload.

AWS DMS supports Amazon S3 as the source and Kinesis as the target, so data stored in an S3 bucket is streamed to Kinesis. Several consumers, such as AWS Lambda, Amazon Kinesis Data Firehose, Amazon Kinesis Data Analytics, and the Kinesis Consumer Library (KCL), can consume the data concurrently to perform real-time analytics on the dataset. Each AWS service in this architecture can scale independently as needed.

You are working as an AWS architect for a weather tracking facility. You are asked to set up a Disaster Recovery (DR) mechanism with minimum costs. In case of failure, the facility can only bear data loss of approximately 15 minutes without jeopardizing the forecasting models.

As a Solutions Architect, which DR method will you suggest?

**Pilot Light**

The term pilot light is often used to describe a DR scenario in which a minimal version of an environment is always running in the cloud. The idea of the pilot light is an analogy that comes from the gas heater. In a gas heater, a small flame that’s always on can quickly ignite the entire furnace to heat up a house. This scenario is similar to a backup-and-restore scenario. For example, with AWS you can maintain a pilot light by configuring and running the most critical core elements of your system in AWS. For the given use-case, a small part of the backup infrastructure is always running simultaneously syncing mutable data (such as databases or documents) so that there is no loss of critical data. When the time comes for recovery, you can rapidly provision a full-scale production environment around the critical core. For Pilot light, RPO/RTO is in 10s of minutes, so this is the correct solution.

An e-commerce company tracks user clicks on its flagship website and performs analytics to provide near-real-time product recommendations. An Amazon EC2 instance receives data from the website and sends the data to an Amazon Aurora Database instance. Another Amazon EC2 instance continuously checks the changes in the database and executes SQL queries to provide recommendations. Now, the company wants a redesign to decouple and scale the infrastructure. The solution must ensure that data can be analyzed in real-time without any data loss even when the company sees huge traffic spikes.

What would you recommend as an AWS Certified Solutions Architect - Associate?

**Leverage Amazon Kinesis Data Streams to capture the data from the website and feed it into Amazon Kinesis Data Analytics which can query the data in real time. Lastly, the analyzed feed is output into Amazon Kinesis Data Firehose to persist the data on Amazon S3**

You can use Amazon Kinesis Data Streams to build custom applications that process or analyze streaming data for specialized needs. Amazon Kinesis Data Streams manages the infrastructure, storage, networking, and configuration needed to stream your data at the level of your data throughput. You don't have to worry about provisioning, deployment, or ongoing maintenance of hardware, software, or other services for your data streams.

For the given use case, you can use Amazon Kinesis Data Analytics to transform and analyze incoming streaming data from Kinesis Data Streams in real time. Kinesis Data Analytics takes care of everything required to run streaming applications continuously, and scales automatically to match the volume and throughput of your incoming data. With Amazon Kinesis Data Analytics, there are no servers to manage, no minimum fee or setup cost, and you only pay for the resources your streaming applications consume.

Question 23: **Incorrect**

The engineering team at a global e-commerce company is currently reviewing their disaster recovery strategy. The team has outlined that they need to be able to quickly recover their application stack with a Recovery Time Objective (RTO) of 5 minutes, in all of the AWS Regions that the application runs. The application stack currently takes over 45 minutes to install on a Linux system.

As a Solutions architect, which of the following options would you recommend as the disaster recovery strategy?

**Create an Amazon Machine Image (AMI) after installing the software and copy the AMI across all Regions. Use this Region-specific AMI to run the recovery process in the respective Regions**

An Amazon Machine Image (AMI) provides the information required to launch an instance. You must specify an AMI when you launch an instance. You can launch multiple instances from a single AMI when you need multiple instances with the same configuration. You can use different AMIs to launch instances when you need instances with different configurations.

For the current use case, you need to create an AMI such that the application stack is already set up. But AMIs are bound to the Region they are created in. So, you need to copy the AMI across Regions for disaster recovery readiness.

Copying a source AMI results in an identical but distinct target AMI with its own unique identifier. In the case of an Amazon EBS-backed AMI, each of its backing snapshots is, by default, copied to an identical but distinct target snapshot. (The sole exceptions are when you choose to encrypt or re-encrypt the snapshot.) You can change or deregister the source AMI with no effect on the target AMI. The reverse is also true. There are no charges for copying an AMI. However, standard storage and data transfer rates apply. If you copy an Amazon EBS-backed AMI, you will incur charges for the storage of any additional Amazon EBS snapshots.

A startup's cloud infrastructure consists of a few Amazon EC2 instances, Amazon RDS instances and Amazon S3 storage. A year into their business operations, the startup is incurring costs that seem too high for their business requirements.

Which of the following options represents a valid cost-optimization solution?

Correct option:

**Use AWS Cost Explorer Resource Optimization to get a report of Amazon EC2 instances that are either idle or have low utilization and use AWS Compute Optimizer to look at instance type recommendations**

AWS Cost Explorer helps you identify under-utilized Amazon EC2 instances that may be downsized on an instance by instance basis within the same instance family, and also understand the potential impact on your AWS bill by taking into account your Reserved Instances and Savings Plans.

AWS Compute Optimizer recommends optimal AWS Compute resources for your workloads to reduce costs and improve performance by using machine learning to analyze historical utilization metrics. Compute Optimizer helps you choose the optimal Amazon EC2 instance types, including those that are part of an Amazon EC2 Auto Scaling group, based on your utilization data.

A systems administrator is creating IAM policies and attaching them to IAM identities. After creating the necessary identity-based policies, the administrator is now creating resource-based policies.

Which is the only resource-based policy that the IAM service supports?

Correct option:

You manage access in AWS by creating policies and attaching them to IAM identities (users, groups of users, or roles) or AWS resources. A policy is an object in AWS that, when associated with an identity or resource, defines their permissions. Resource-based policies are JSON policy documents that you attach to a resource such as an Amazon S3 bucket. These policies grant the specified principal permission to perform specific actions on that resource and define under what conditions this applies.

**Trust policy**

Trust policies define which principal entities (accounts, users, roles, and federated users) can assume the role. An IAM role is both an identity and a resource that supports resource-based policies. For this reason, you must attach both a trust policy and an identity-based policy to an IAM role. The IAM service supports only one type of resource-based policy called a role trust policy, which is attached to an IAM role.

A retail company uses AWS Cloud to manage its technology infrastructure. The company has deployed its consumer-focused web application on Amazon EC2-based web servers and uses Amazon RDS PostgreSQL database as the data store. The PostgreSQL database is set up in a private subnet that allows inbound traffic from selected Amazon EC2 instances. The database also uses AWS Key Management Service (AWS KMS) for encrypting data at rest.

Which of the following steps would you recommend to facilitate end-to-end secure access to the data in the database?

**Configure Amazon RDS to use SSL for data in transit**

You can use Secure Socket Layer / Transport Layer Security (SSL/TLS) connections to encrypt data in transit. Amazon RDS creates an SSL certificate and installs the certificate on the DB instance when the instance is provisioned. For MySQL, you launch the MySQL client using the --ssl\_ca parameter to reference the public key to encrypt connections. Using SSL, you can encrypt a PostgreSQL connection between your applications and your PostgreSQL DB instances. You can also force all connections to your PostgreSQL DB instance to use SSL.

Your company is deploying a website running on AWS Elastic Beanstalk. The website takes over 45 minutes for the installation and contains both static as well as dynamic files that must be generated during the installation process.

As a Solutions Architect, you would like to bring the time to create a new instance in your AWS Elastic Beanstalk deployment to be less than 2 minutes. Which of the following options should be combined to build a solution for this requirement? (Select two)

**Create a Golden Amazon Machine Image (AMI) with the static installation components already setup**

A Golden AMI is an AMI that you standardize through configuration, consistent security patching, and hardening. It also contains agents you approve for logging, security, performance monitoring, etc. For the given use-case, you can have the static installation components already setup via the golden AMI.

**Use Amazon EC2 user data to customize the dynamic installation parts at boot time**

Amazon EC2 instance user data is the data that you specified in the form of a configuration script while launching your instance. You can use Amazon EC2 user data to customize the dynamic installation parts at boot time, rather than installing the application itself at boot time.

A media company uses Amazon ElastiCache Redis to enhance the performance of its Amazon RDS database layer. The company wants a robust disaster recovery strategy for its caching layer that guarantees minimal downtime as well as minimal data loss while ensuring good application performance.

Which of the following solutions will you recommend to address the given use-case?

**Opt for Multi-AZ configuration with automatic failover functionality to help mitigate failure**

Multi-AZ is the best option when data retention, minimal downtime, and application performance are a priority.

Data-loss potential - Low. Multi-AZ provides fault tolerance for every scenario, including hardware-related issues.

Performance impact - Low. Of the available options, Multi-AZ provides the fastest time to recovery, because there is no manual procedure to follow after the process is implemented.

Cost - Low to high. Multi-AZ is the lowest-cost option. Use Multi-AZ when you can't risk losing data because of hardware failure or you can't afford the downtime required by other options in your response to an outage.

As an e-sport tournament hosting company, you have servers that need to scale and be highly available. Therefore you have deployed an Elastic Load Balancing (ELB) with an Auto Scaling group (ASG) across 3 Availability Zones (AZs). When e-sport tournaments are running, the servers need to scale quickly. And when tournaments are done, the servers can be idle. As a general rule, you would like to be highly available, have the capacity to scale and optimize your costs.

What do you recommend? (Select two)

**Set the minimum capacity to 2**

An Auto Scaling group contains a collection of Amazon EC2 instances that are treated as a logical grouping for automatic scaling and management. An Auto Scaling group also enables you to use Amazon EC2 Auto Scaling features such as health check replacements and scaling policies. Both maintaining the number of instances in an Auto Scaling group and automatic scaling are the core functionality of the Amazon EC2 Auto Scaling service.

You configure the size of your Auto Scaling group by setting the minimum, maximum, and desired capacity. The minimum and maximum capacity are required to create an Auto Scaling group, while the desired capacity is optional. If you do not define your desired capacity upfront, it defaults to your minimum capacity.

An Auto Scaling group is elastic as long as it has different values for minimum and maximum capacity. All requests to change the Auto Scaling group's desired capacity (either by manual scaling or automatic scaling) must fall within these limits.

Here, even though our ASG is deployed across 3 Availability Zones (AZs), the minimum capacity to be highly available is 2. When we specify 2 as the minimum capacity, the ASG would create these 2 instances in separate Availability Zones (AZs). If demand goes up, the ASG would spin up a new instance in the third Availability Zone (AZ). Later as the demand subsides, the ASG would scale-in and the instance count would be back to 2.

**Use Reserved Instances (RIs) for the minimum capacity**

Reserved Instances (RIs) provide you with significant savings on your Amazon EC2 costs compared to On-Demand Instance pricing. Reserved Instances are not physical instances, but rather a billing discount applied to the use of On-Demand Instances in your account. These On-Demand Instances must match certain attributes, such as instance type and Region, to benefit from the billing discount. Since minimum capacity will always be maintained, it is cost-effective to choose reserved instances than any other option.

In case of an Availability Zone (AZ) outage, the instance in that Availability Zone (AZ) would go down however the other instance would still be available. The ASG would provision the replacement instance in the third Availability Zone (AZ) to keep the minimum count to 2.

What does this AWS CloudFormation snippet do? (Select three)

**SecurityGroupIngress:**

**- IpProtocol: tcp**

**FromPort: 80**

**ToPort: 80**

**CidrIp: 0.0.0.0/0**

**- IpProtocol: tcp**

**FromPort: 22**

**ToPort: 22**

**CidrIp: 192.168.1.1/32**

**It allows any IP to pass through on the HTTP port**

**It configures a security group's inbound rules**

**It lets traffic flow from one IP on port 22**

A security group acts as a virtual firewall that controls the traffic for one or more instances. When you launch an instance, you can specify one or more security groups; otherwise, we use the default security group. You can add rules to each security group that allows traffic to or from its associated instances. You can modify the rules for a security group at any time; the new rules are automatically applied to all instances that are associated with the security group. When we decide whether to allow traffic to reach an instance, we evaluate all the rules from all the security groups that are associated with the instance.

The following are the characteristics of security group rules: 1. By default, security groups allow all outbound traffic. 2. Security group rules are always permissive; you can't create rules that deny access. 3. Security groups are stateful

An IT company has a large number of clients opting to build their application programming interface (API) using Docker containers. To facilitate the hosting of these containers, the company is looking at various orchestration services available with AWS.

As a Solutions Architect, which of the following solutions will you suggest? (Select two)

Correct options:

**Use Amazon Elastic Kubernetes Service (Amazon EKS) with AWS Fargate for serverless orchestration of the containerized services**

**Use Amazon Elastic Container Service (Amazon ECS) with AWS Fargate for serverless orchestration of the containerized services**

Building APIs with Docker containers has been gaining momentum over the years. For hosting and exposing these container-based APIs, they need a solution which supports HTTP requests routing, autoscaling, and high availability. In some cases, user authorization is also needed.

For this purpose, many organizations are orchestrating their containerized services with Amazon Elastic Container Service (Amazon ECS) or Amazon Elastic Kubernetes Service (Amazon EKS), while hosting their containers on Amazon EC2 or AWS Fargate. Then, they can add scalability and high availability with Service Auto Scaling (in Amazon ECS) or Horizontal Pod Auto Scaler (in Amazon EKS), and they expose the services through load balancers.

When you use Amazon ECS as an orchestrator (with EC2 or Fargate launch type), you also have the option to expose your services with Amazon API Gateway and AWS Cloud Map instead of a load balancer. AWS Cloud Map is used for service discovery: no matter how Amazon ECS tasks scale, AWS Cloud Map service names would point to the right set of Amazon ECS tasks. Then, API Gateway HTTP APIs can be used to define API routes and point them to the corresponding AWS Cloud Map services.

A social media company wants the capability to dynamically alter the size of a geographic area from which traffic is routed to a specific server resource.

Which feature of Amazon Route 53 can help achieve this functionality?

* **Geoproximity routing**

**(Correct)**

* **Geolocation routing**

**(Incorrect)**

* **Latency-based routing**
* **Weighted routing**
* A small rental company had 5 employees, all working under the same AWS cloud account. These employees deployed their applications built for various functions- including billing, operations, finance, etc. Each of these employees has been operating in their own VPC. Now, there is a need to connect these VPCs so that the applications can communicate with each other.
* Which of the following is the MOST cost-effective solution for this use-case?
* **Use a VPC peering connection**
* A VPC peering connection is a networking connection between two VPCs that enables you to route traffic between them using private IPv4 addresses or IPv6 addresses. Instances in either VPC can communicate with each other as if they are within the same network. You can create a VPC peering connection between your own VPCs, or with a VPC in another AWS account. The VPCs can be in different regions (also known as an inter-region VPC peering connection). VPC Peering helps connect two VPCs and is not transitive. To connect VPCs together, the best available option is to use VPC peering.
* A junior developer has downloaded a sample Amazon S3 bucket policy to make changes to it based on new company-wide access policies. He has requested your help in understanding this bucket policy.
* As a Solutions Architect, which of the following would you identify as the correct description for the given policy?
* **{**
* **"Version": "2012-10-17",**
* **"Id": "S3PolicyId1",**
* **"Statement": [**
* **{**
* **"Sid": "IPAllow",**
* **"Effect": "Allow",**
* **"Principal": "\*",**
* **"Action": "s3:\*",**
* **"Resource": "arn:aws:s3:::examplebucket/\*",**
* **"Condition": {**
* **"IpAddress": {"aws:SourceIp": "54.240.143.0/24"},**
* **"NotIpAddress": {"aws:SourceIp": "54.240.143.188/32"}**
* **}**
* **}**
* **]**
* **}**
* You manage access in AWS by creating policies and attaching them to IAM identities (users, groups of users, or roles) or AWS resources. A policy is an object in AWS that, when associated with an identity or resource, defines their permissions. AWS evaluates these policies when an IAM principal (user or role) makes a request. Permissions in the policies determine whether the request is allowed or denied. Most policies are stored in AWS as JSON documents. AWS supports six types of policies: identity-based policies, resource-based policies, permissions boundaries, AWS Organizations SCPs, ACLs, and session policies.
* Let's analyze the bucket policy one step at a time:
* The snippet "Effect": "Allow" implies an allow effect. The snippet "Principal": "\*" implies any Principal. The snippet "Action": "s3:\*" implies any Amazon S3 API. The snippet "Resource": "arn:aws:s3:::examplebucket/\*" implies that the resource can be the bucket examplebucket and its contents. Consider the last snippet of the given bucket policy: "Condition": { "IpAddress": {"aws:SourceIp": "54.240.143.0/24"}, "NotIpAddress": {"aws:SourceIp": "54.240.143.188/32"} } This snippet implies that if the source IP is in the CIDR block "54.240.143.0/24" (== 54.240.143.0 - 54.240.143.255), then it is allowed to access the examplebucket and its contents. However, the source IP cannot be in the CIDR "54.240.143.188/32" (== 1 IP, 54.240.143.188/32), which means one IP address is explicitly blocked from accessing the examplebucket and its contents.

Your company runs a web portal to match developers to clients who need their help. As a solutions architect, you've designed the architecture of the website to be fully serverless with Amazon API Gateway and AWS Lambda. The backend uses Amazon DynamoDB table. You would like to automatically congratulate your developers on important milestones, such as - their first paid contract. All the contracts are stored in Amazon DynamoDB.

Which Amazon DynamoDB feature can you use to implement this functionality such that there is LEAST delay in sending automatic notifications?

* **Amazon DynamoDB DAX + Amazon API Gateway**

**(Incorrect)**

* **Amazon Simple Queue Service (Amazon SQS) + AWS Lambda**
* **Amazon EventBridge events + AWS Lambda**
* **Amazon DynamoDB Streams + AWS Lambda**

**(Correct)**

A company uses Application Load Balancers in multiple AWS Regions. The Application Load Balancers receive inconsistent traffic that varies throughout the year. The engineering team at the company needs to allow the IP addresses of the Application Load Balancers in the on-premises firewall to enable connectivity.

Which of the following represents the MOST scalable solution with minimal configuration changes?

Correct option:

**Set up AWS Global Accelerator. Register the Application Load Balancers in different Regions to the AWS Global Accelerator. Configure the on-premises firewall's rule to allow static IP addresses associated with the AWS Global Accelerator**

AWS Global Accelerator is a networking service that helps you improve the availability and performance of the applications that you offer to your global users. AWS Global Accelerator is easy to set up, configure, and manage. It provides static IP addresses that provide a fixed entry point to your applications and eliminate the complexity of managing specific IP addresses for different AWS Regions and Availability Zones.

Associate the static IP addresses provided by AWS Global Accelerator to regional AWS resources or endpoints, such as Network Load Balancers, Application Load Balancers, Amazon EC2 Instances, and Elastic IP addresses. The IP addresses are anycast from AWS edge locations so they provide onboarding to the AWS global network close to your users.

An e-commerce company has copied 1 petabyte of data from its on-premises data center to an Amazon S3 bucket in the us-west-1 Region using an AWS Direct Connect link. The company now wants to set up a one-time copy of the data to another Amazon S3 bucket in the us-east-1 Region. The on-premises data center does not allow the use of AWS Snowball.

As a Solutions Architect, which of the following options can be used to accomplish this goal? (Select two)

**Copy data from the source bucket to the destination bucket using the aws S3 sync command**

The aws S3 sync command uses the CopyObject APIs to copy objects between Amazon S3 buckets. The sync command lists the source and target buckets to identify objects that are in the source bucket but that aren't in the target bucket. The command also identifies objects in the source bucket that have different LastModified dates than the objects that are in the target bucket. The sync command on a versioned bucket copies only the current version of the object—previous versions aren't copied. By default, this preserves object metadata, but the access control lists (ACLs) are set to FULL\_CONTROL for your AWS account, which removes any additional ACLs. If the operation fails, you can run the sync command again without duplicating previously copied objects.

You can use the command like so:

aws s3 sync s3://DOC-EXAMPLE-BUCKET-SOURCE s3://DOC-EXAMPLE-BUCKET-TARGET

**Set up Amazon S3 batch replication to copy objects across Amazon S3 buckets in another Region using S3 console and then delete the replication configuration**

Amazon S3 Batch Replication provides you a way to replicate objects that existed before a replication configuration was in place, objects that have previously been replicated, and objects that have failed replication. This is done through the use of a Batch Operations job.

You should note that batch replication differs from live replication which continuously and automatically replicates new objects across Amazon S3 buckets. You cannot directly use the AWS S3 console to configure cross-Region replication for existing objects. By default, replication only supports copying new Amazon S3 objects after it is enabled using the AWS S3 console. Replication enables automatic, asynchronous copying of objects across Amazon S3 buckets. Buckets that are configured for object replication can be owned by the same AWS account or by different accounts. Object may be replicated to a single destination bucket or multiple destination buckets. Destination buckets can be in different AWS Regions or within the same Region as the source bucket. Once done, you can delete the replication configuration, as it ensures that batch replication is only used for this one-time data copy operation.

If you want to enable live replication for existing objects for your bucket, you must contact AWS Support and raise a support ticket. This is required to ensure that replication is configured correctly.

A CRM company has a software as a service (SaaS) application that feeds updates to other in-house and third-party applications. The SaaS application and the in-house applications are being migrated to use AWS services for this inter-application communication.

As a Solutions Architect, which of the following would you suggest to asynchronously decouple the architecture?

* **Use Amazon EventBridge to decouple the system architecture**

**(Correct)**

* **Use Elastic Load Balancing (ELB) for effective decoupling of system architecture**
* **Use Amazon Simple Notification Service (Amazon SNS) to communicate between systems and decouple the architecture**
* **Use Amazon Simple Queue Service (Amazon SQS) to decouple the architecture**

**(Incorrect)**

The engineering team at a social media company has recently migrated to AWS Cloud from its on-premises data center. The team is evaluating Amazon CloudFront to be used as a CDN for its flagship application. The team has hired you as an AWS Certified Solutions Architect – Associate to advise on Amazon CloudFront capabilities on routing, security, and high availability.

Which of the following would you identify as correct regarding Amazon CloudFront? (Select three)

**Amazon CloudFront can route to multiple origins based on the content type**

You can configure a single Amazon CloudFront web distribution to serve different types of requests from multiple origins. For example, if you are building a website that serves static content from an Amazon Simple Storage Service (Amazon S3) bucket and dynamic content from a load balancer, you can serve both types of content from a Amazon CloudFront web distribution.

**Use an origin group with primary and secondary origins to configure Amazon CloudFront for high-availability and failover**

You can set up Amazon CloudFront with origin failover for scenarios that require high availability. To get started, you create an origin group with two origins: a primary and a secondary. If the primary origin is unavailable or returns specific HTTP response status codes that indicate a failure, CloudFront automatically switches to the secondary origin.

To set up origin failover, you must have a distribution with at least two origins. Next, you create an origin group for your distribution that includes two origins, setting one as the primary. Finally, you create or update a cache behavior to use the origin group.

The development team at a social media company wants to handle some complicated queries such as "What are the number of likes on the videos that have been posted by friends of a user A?".

As a solutions architect, which of the following AWS database services would you suggest as the BEST fit to handle such use cases?

**Amazon Neptune**

Amazon Neptune is a fast, reliable, fully managed graph database service that makes it easy to build and run applications that work with highly connected datasets. The core of Amazon Neptune is a purpose-built, high-performance graph database engine optimized for storing billions of relationships and querying the graph with milliseconds latency. Neptune powers graph use cases such as recommendation engines, fraud detection, knowledge graphs, drug discovery, and network security.

Amazon Neptune is highly available, with read replicas, point-in-time recovery, continuous backup to Amazon S3, and replication across Availability Zones. Neptune is secure with support for HTTPS encrypted client connections and encryption at rest. Neptune is fully managed, so you no longer need to worry about database management tasks such as hardware provisioning, software patching, setup, configuration, or backups.

Amazon Neptune can quickly and easily process large sets of user-profiles and interactions to build social networking applications. Neptune enables highly interactive graph queries with high throughput to bring social features into your applications. For example, if you are building a social feed into your application, you can use Neptune to provide results that prioritize showing your users the latest updates from their family, from friends whose updates they ‘Like,’ and from friends who live close to them.